Exploring patterns of variation in clutch size-density reaction norms in a wild passerine bird

Nicolaus, M.; Brommer, J. E.; Ubels, R.; Tinbergen, J. M.; Dingemanse, N. J.

Published in:
Journal of Evolutionary Biology

DOI:
10.1111/jeb.12210

Publication date:
2013

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12210

Exploring patterns of variation in clutch size–density reaction norms in a wild passerine bird

M. NICOLAUS*, J. E. BROMMER†, R. UBELS‡, J. M. TINBERGEN‡ & N. J. DINGEMANSE*§

*Evolutionary Ecology of Variation Group, Max Planck Institute for Ornithology, Seewiesen, Germany
†Bird Ecology Unit, Department of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland
‡Animal Ecology Group, Centre for Ecological and Evolutionary Studies, University of Groningen, CC Groningen, The Netherlands
§Department Biologie II, Behavioural Ecology Group, Ludwig Maximilians University of Munich, Planegg-Martinsried, Germany

Keywords:

animal model;
density dependence;
heritability;
heteroscedasticity;
Parus major;
phenotypic plasticity;
random regression;
reaction norm;
residual variance.

Abstract

Negative density dependence of clutch size is a ubiquitous characteristic of avian populations and is partly due to within-individual phenotypic plasticity. Yet, very little is known about the extent to which individuals differ in their degree of phenotypic plasticity, whether such variation has a genetic basis and whether level of plasticity can thus evolve in response to selection. Using 18 years of data of a Dutch great tit population (*Parus major*), we show that females reduced clutch size with increasing population density (slopes of the reaction norms), differed strongly in their average clutch size (elevations of the reaction norms) at the population-mean density and that the latter variation was partly heritable. In contrast, we could not detect individual variation in phenotypic plasticity ('I × E'). Level of plasticity is thus not likely to evolve in response to selection in this population. Observed clutch sizes deviated more from the estimated individual reaction norms in certain years and densities, implying that the within-individual between-year variance (so-called residual variance) of clutch size was heterogeneous with respect to these factors. Given the observational nature of this study, experimental manipulation of density is now warranted to confirm the causality of the observed density effects. Our analyses demonstrate that failure to acknowledge this heterogeneity would have inflated the estimate of 'I × E' and led to misinterpretation of the data. This paper thereby emphasizes the fact that heterogeneity in residuals can provide biologically insightful information about the ecological processes underlying the data.

Introduction

Phenotypic plasticity is a ubiquitous feature of animal populations and characterizes the ability of a single individual (or genotype) to alter its phenotype in response to changes in environmental conditions (DeWitt & Scheiner, 1994). Behavioural ecologists have long been interested in this phenomenon as plasticity represents an important mechanism by which individuals or populations can respond adaptively to rapidly changing conditions (Piersma & Drent, 2003) such as climate change (Nussey *et al.*, 2005b; Brommer *et al.*, 2008; Charmantier *et al.*, 2008) or changes in perceived predation risk (e.g. Martin & Réale, 2008; Lima, 2009). Over the last few years, ecologists have increasingly used repeated measures data from longitudinal studies to quantify variation in plasticity at different levels, for example, among populations and among individuals within the same population. These studies have revealed that individuals within the same population often differ in their degree of plasticity (behaviours: review in Dingemanse *et al.*, 2010; Mathot *et al.*, 2012; Brommer, 2013; and life-history traits: review in Nussey *et al.*, 2007). Individual variation in plasticity

Correspondence: Marion Nicolaus, Evolutionary Ecology of Variation Group, Max Planck Institute for Ornithology, E-Gwinner Str. 7, 82319 Seewiesen, Germany. Tel.: +49 (0)8157 932 231; fax: +49 (0)8157 932 209; e-mail: mnicolaus@orn.mpg.de

may be important for evolutionary processes because if it has a genetic basis, it has the potential to evolve under natural selection and provide important information on how quickly natural populations may respond to changes in their environment (Nussey *et al.*, 2007). Recently, studies in behavioural ecology have also emphasized that selection might favour the evolution of between-individual variation in plasticity (e.g. Wolf *et al.*, 2008, 2011), and therefore, the study of heterogeneity in phenotypic plasticity – for long the realm of evolutionary biologists – has now entered the field of behavioural ecology (see also van de Pol, 2012; Dingemanse & Wolf, 2013).

Phenotypic adjustment of labile traits is typically conceptualized and measured in terms of reaction norms where the (genotype's) phenotypic response is regressed against an environmental gradient (Via *et al.*, 1995). In this study, we study individual variation in clutch size–density reaction norms in a wild great tit (*Parus major*) population with known pedigree structure. Density dependence of reproduction is a key concept in ecology because of its role in regulating population sizes (Newton, 1998). In birds, longitudinal studies have revealed that individuals typically produce fewer offspring in high-density years (Both, 2000) and that this pattern is often the result of within-individual plasticity rather than selective (dis) appearance of phenotypes (Both, 1998a; Nicolaus *et al.*, 2009a). It has been hypothesized that individual adjustment of clutch size to population density is an adaptive response to an increase in competition for local resources (Both *et al.*, 2000) or is constrained by limited resources (Arcese & Smith, 1988; Nager *et al.*, 1997). This suggests that the clutch size–density reaction norms may be under selection because individuals that do not adjust their clutch size to density should suffer reduced fitness. Yet, despite a large body of studies documenting density dependence of avian reproduction, surprisingly little is known about the extent to which individuals differ in their adjustment of clutch size to the prevailing density ('I × E'), and whether such variation has a genetic basis ('G × E'). This information is nevertheless of importance because it would enrich our understanding of whether there is indeed individual-level adjustment of reproduction to population density and whether such a phenotypic response can evolve.

Here, we therefore examine two central aspects of the clutch size–density relationship using 18 years of data of a pedigreed population. First, we confirm that females reduce clutch size with increasing population density and then test whether females do differ in the slope of their clutch size–density reaction norm (I × E). Second, we test whether variation in plasticity in the clutch size–density relationship has a genetic basis (G × E), that is, whether it can evolve under selection.

Materials and methods

Study area and study species

The study was carried out in a great tit population inhabiting a mixed deciduous forest in the Lauwersmeer area in the north of the Netherlands. This population has been monitored since 1993; the number and position of the nest boxes in the study have changed according to the experiments performed over the study years (see details in Tinbergen & Sanz, 2004; Nicolaus *et al.*, 2009b): First, from 1994 onwards, 200 nest boxes were unevenly spread over eight nest-box areas ('plots'), where each plot had either a high (from 1.25 to 1.64 nest boxes ha^{-1}) or low (from 0.11 to 0.83 nest boxes ha^{-1}) number of boxes (phase 1). Second, in 2004, 200 nest boxes were added in 4 additional plots, offering in total 400 available nest sites (phase 2). Third, before the 2005 breeding season, we completely reorganized the existing study area by establishing 12 new plots. From that point onwards, each plot consisted of 50 boxes totalling 600 nest sites available for breeding (phase 3). Nest boxes were spread over nine plots in phase 1, 13 plots in phase 2 and 12 plots in phase 3. In this study, we distinguish two periods characterized by overall 'low' (period 1, 1994–2004) or 'high' (period 2, 2005–2011) population density (Fig. 1a). Population density significantly increased over the years in period 1 ($r_s = 0.87$, $P = 0.002$, $n = 11$ years) but not in period 2 ($r_s = -0.05$, $P = 0.935$, $n = 7$ years). Given the relatively young age of the forest (established between 1971 and 1988) and the low density of excavator birds (i.e. great-spotted woodpeckers *Dendrocopos major*), very few natural cavities were available in the study area, that is, the vast majority of breeding attempts occurred in the nest boxes (Nicolaus *et al.*, 2009a).

Data collection

From the beginning of April, nest boxes were checked weekly and parameters such as lay date, clutch size, parental identity, nestling growth and fledging success were monitored using standard methods (detailed in Nicolaus *et al.*, 2009b). Minimum age of captured parents was inferred from either previous data (i.e. for recruited birds that were ringed as nestlings in our populations) or plumage characteristics (i.e. for unbanded immigrant birds; Svensson, 1992). Further weekly checks gave information about the incidence of second clutches and their breeding success.

Data selection

We used 18 years of breeding data (1994–2011), focussing on first clutches only. First clutches were defined as clutches that were initiated within 30 days of the

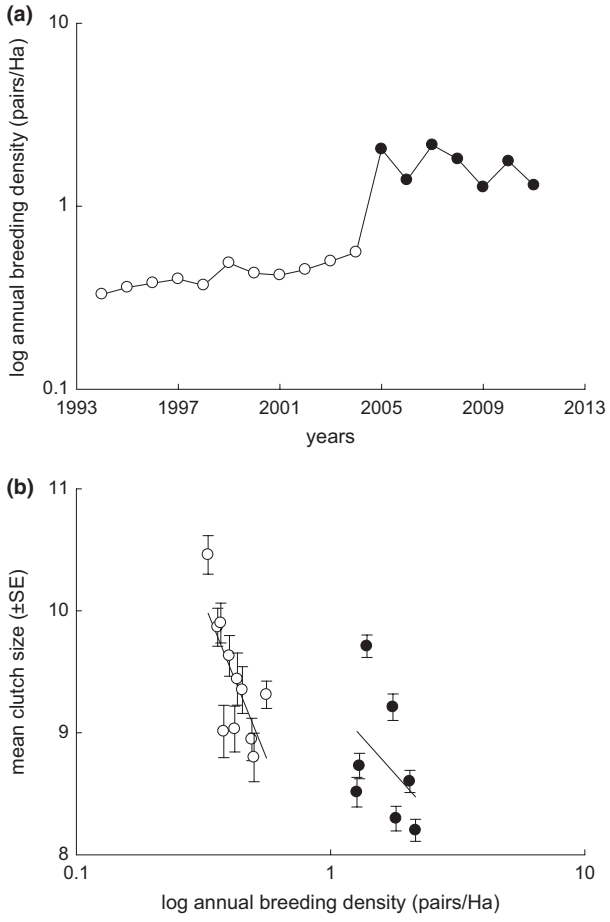

Fig. 1 (a) Breeding density of great tits has increased over time (n = 18 years) in the Lauwersmeer (the Netherlands). (b) Mean population clutch size (±standard error) decreased with breeding density (n = 2464 observations of 1895 females, raw data). 'Low'- and 'high'-density periods are denoted in white and black, respectively.

start of the earliest clutch in the focal year. Clutches were excluded from the analyses when the female identity was unknown or when clutches were smaller than five eggs or larger than 15 eggs because these are often the result of disturbed situations (i.e. clutch desertion or multiple females laying in one nest, n = 155 events). Our data set includes on average (SE) 1.44 ± 0.02 breeding events per female (n = 1895). The number of breeding events per female ranges from 1 to 6 with a mode and a median of 1 (1195 female with one observation, 351 females with two observations; 110 females with three observations; 40 females with four observations; 13 females with five observations and two females with six observations). These data constitute a substantial data set for exploring individual variation in reaction norms (see power analysis in Appendix S1). We note that within-individual variance in clutch size here represents among-year within-individual variance.

Breeding densities

Breeding density was defined as the number of breeding pairs per surface area (pairs ha^{-1}). We have previously shown that females adjust their clutch size to the population annual breeding density (=number of annual breeding pairs per total surface area; varying among years) and not to local plot breeding densities (=number of annual breeding pairs per plot surface area; varying among plots within and among years). This pattern was true even when we experimentally induced a sevenfold difference in local breeding densities between the plots in period 1 (Nicolaus *et al.*, 2009a). In period 2, natural variation in local plot breeding densities was small because all plots were provided with the same number of nest boxes (natural difference of maximum 2.6-fold on average) and thus unlikely to influence female reproductive decisions. We therefore used population annual breeding density as the environmental covariate in the reaction norms. Changes in the size of the study area across years were taken into account in the calculations of density by dividing the number of breeding birds by the size of the study area for each year separately (Fig. 1a,b).

Random regression approaches

In a reaction norm approach, the phenotypic response of one individual or genotype to the environment is characterized by an elevation (its expected trait value in the average environment, 'I') and a slope (amount by which the trait changes per unit change of the environment, 'E', that is, phenotypic plasticity; Pigliucci, 2005). These two reaction norm components can themselves be considered as quantitative traits, where random regression phenotypic model (RRPM; Henderson, 1982) can be applied to test for the presence of individual variation in plasticity (i.e. individual by environment interaction, 'I × E'; Nussey *et al.*, 2007) and correlations between elevation and slope. When pedigree data are available, random regression animal models (RRAM) are used to further quantify the relative contribution of genetic and nongenetic sources of variation in both reaction norm components (Meyer, 1998; Kruuk, 2004; Schaeffer, 2004; Wilson *et al.*, 2010). RRAM thus provides a means to partition 'I' into 'G' (effects of genes) and 'PE' (permanent effects of environments, that is, individual-specific nonheritable effects) and to partition 'I × E' into its heritable component ('G × E', genotype by environment interaction) and nonheritable component ('PE × E'). The permanent environment effect ('PE') captures the influence of environmental factors that are associated with an individual and whose effects are conserved across the

repeated records of that individual. This environmental effect is, in any case, different from the within-individual variance (or 'residual variance').

Random regression phenotypic model

The clutch size $CS_{i,i,k}$ of individual i in year j and plot k was modelled as a continuous function of the year-specific breeding density D. We used 2784 measures of 1895 individuals, including those with only one measurement (Martin *et al.*, 2011). With such a sample size, the power to detect individual differences in elevations at intermediate population size for the Lauwersmeer population (i.e. $N = 175$ females sampled) is around 1, and the power to detect individual differences in slopes ranges between 0.60 and 1 (see power analysis in Appendix S1; van de Pol, 2012). Following the procedure described by Brommer *et al.* (2012), we first modelled the clutch size–density relationships at the phenotypic level using a RRPM:

$$\begin{aligned} CS_{i,j,k} = {} & \beta_0 + \beta_1 LD_{i,j} + \beta_2 Age_{i,j,k} + \beta_3 Age^2_{i,j,k} + \beta_4 D_j \\ & + year_j + plot_k + f(ind^{CS}_{x,i}, D_j) + \varepsilon_{i,j,k,l} \end{aligned} \quad (1)$$

where eqn (1) specifies a mixed-effect model with β_0 (the overall mean clutch size), β_1 LD, β_2 Age, β_3 Age^2 and β_4 D fitted as fixed effects and year, plot and $f(ind^{CS}_{x,i}, D_j)$ as random effects, and $\varepsilon_{i,j,k,l}$ denoting the residual error. We corrected for parameters known to influence clutch size: $LD_{i,j}$, the laying date of individual i in year j (Klomp, 1970), $Age_{i,j,k}$ and $Age^2_{i,j,k}$ the age of individual i and its quadrate in year j to account for nonlinear effects of age (Kluijver, 1951; Bouwhuis *et al.*, 2010). D_j is the year-specific breeding density standardized to zero mean and standard deviation unit, and 'β_4' is the fixed-effect regression coefficient of the mean clutch size changing with D. Individual variation that is not explained by the fixed effects is modelled by the random effects and residuals, where the random effects *year* and *plot* capture annual and spatial variation. Because *year* was fitted as a random effect into the model, the effect of density was consequently not pseudo-replicated. The polynomial function $f(ind^{CS}_{x,i}, D_j)$ specifies the individual-specific deviations from the fixed-effect slope of clutch size on breeding density. For each individual i, a polynomial function of increasing order x is specified. When $x = 0$, the variance across individuals ind_0 (elevation) is estimated at the average density value (i.e. is stable over the density gradient). When $x = 1$, the variance in the coefficients ind_0 (elevation) and the variance in ind_1 (slope) of the function $ind_0 + ind_1$ are estimated as well as the covariance between them. The residual variance $\varepsilon_{i,j,k,l}$ describes for individual i in year j and plot k at the instance l the deviation from the model's predicted value. When $x = 1$, eqn 1 thus partitions the total phenotypic variance (V_P) into year (V_{YEAR}), plot (V_{PLOT}), between-individual variation in elevation ($V_{ind0,}$ or 'I'), between-individual variation in slope ($V_{ind1,}$ or 'I × E'), covariance between elevation and slope ($COV_{ind0,\ ind1}$) and residual (within-individual) variance (V_R or 'ε'). Polynomial functions where $x = 2$ were also considered but caused model conversion problems and therefore these models will not be reported here.

Random regression animal model

We used restricted maximum-likelihood ('REML') models with a pedigree based on social matings to estimate variance components of clutch size. Such RRAM allow for the decomposition of the phenotypic variance into random and fixed-effect variance components by comparing phenotypes of known relatives using pedigree information and facilitate quantitative genetic analyses of unbalanced data sets (Lynch & Walsh, 1998; Kruuk, 2004). We ran the animal models using a fixed-effect structure detailed in eqn 1 for a pedigree that includes, when pruned to retain only links to the individuals with measurements of clutch size (Morrissey & Wilson, 2010), 2565 records with 1283 maternities and paternities. Maximum pedigree depth was 15 generations. Inbreeding was low (only 83 individuals had a nonzero inbreeding coefficient). Mean maternal sibship size was 1.72 and mean paternal sibship size was 1.66, with a pedigree-wide mean pairwise relatedness of 0.00152. The social pedigree was built using all information about individuals marked in our nest boxes between 1993 and 2011 and based on the assumptions that (i) immigrants are unrelated to each other and to resident birds, (ii) the social parents are also the genetic parents, and (iii) the occurrence of extra-pair paternity (about 10% for this population; Brommer *et al.*, 2010b) has little effect on the estimate of heritability (Charmantier & Réale, 2005). Following procedures described by Nussey *et al.* (2007) and Wilson *et al.* (2010), we expanded eqn 1 by splitting up the phenotypic variance in both elevation and slope into their additive genetic vs. permanent environment components, denoted by the functions $f(a^{CS}_{x,i}, D)$ *and* $f(pe^{CS}_{x,i}, D)$, respectively:

$$\begin{aligned} CS_{i,j,k} = {} & \beta_0 + \beta_1 LD_{i,j} + \beta_2 Age_{i,j,k} + \beta_3 Age^2_{i,j,k} + \beta_4 D_j \\ & + year_j + plot_k + f(a^{CS}_{x,i}, D_j) + f(pe^{CS}_{x,i}, D_j) + \varepsilon_{i,j,k,l} \end{aligned} \quad (2)$$

When $x = 1$, eqn 2 thus partitions the total phenotypic variance (V_P) into year (V_{YEAR}) and plot (V_{PLOT}) variance and into additive genetic effects on elevation (V_{A0} or G) and on slope (V_{A1} or G × E) and their covariance ($COV_{A0,\ A1}$) and permanent environment effects on elevation (V_{PE0} or PE) and on slope (V_{PE1} or PE × E) and their covariance ($COV_{PE0,\ PE1}$).

The statistical significance of all random effects was derived from likelihood ratio tests (LRTs). The test statistic is twice the difference in log-likelihood between

hierarchical nested models and is distributed as χ^2 (Shaw, 1991). Because a (co)variance matrix must be positive definite, the values of variances are bounded to be positive and the values of covariances are also constrained to a range where correlations are between −1 and +1. For this reason, we calculated the probability to obtain a certain χ^2 value while assuming a mixture of χ^2 distributions with differing degrees of freedom (cf. Self & Liang, 1987; Pinheiro & Bates, 2000; Visscher, 2006). Probability (P) of the LRT of a variance was against a mixture of P(χ^2, d.f. = 0) and P(χ^2, d.f. = 1), and the probability of LRT of one variance and one covariance (i.e. the inclusion of variance in reaction norm slope) was tested assuming an equal mixture of P (χ^2, d.f. = 1) and P(χ^2, d.f. = 2). Our findings are qualitatively the same when applying the more conservative testing approach and assuming the d.f. of a LRT equals the difference in the number of (co)variances estimated between two nested models (Results not shown). The significance of the fixed effects was derived using numerator and denumerator degrees of freedom estimated from the algebraic algorithm in ASREML 3.0 (Gilmour *et al.*, 2009). To circumvent the potential problem of producing overconfident estimates of fixed effects, particularly for population density (Schielzeth & Forstmeier, 2009), we reported fixed-effect estimates from the most parsimonious models and from models fitted with random slopes (I × E). All models were implemented in ASREML 3.0 (Gilmour *et al.*, 2009).

Homogeneous vs. heterogeneous residual errors

Typically, random regression models used to test for the presence of individual variation in plasticity ('I × E') assume that the within-individual variance (statistically called 'residual variance') is homogeneous across the environmental covariate 'E', and therefore heterogeneity in residuals is often treated as a statistical nuisance (Erceg-Hurn & Mirosevich, 2008). Although this topic is still in its infancy, an increasing number of studies emphasize the fact that heterogeneity in residuals can provide biologically insightful information about the ecological processes underlying the data (Brommer *et al.*, 2010a; Réale & Dingemanse, 2010; Cleasby & Nakagawa, 2011; Stamps *et al.*, 2012; Westneat *et al.*, 2013). For example, a decrease in phenotypic trait variance along an environmental quality gradient may occur because individuals converge in their responses under adverse conditions (i.e. a reduction in the between-individual variance; Fig. 2b) or because the within-individual (residual) variance decreases under adverse conditions (heterogeneous residual variance; Fig. 2a; Schaeffer, 2004). Hence, ignoring heterogeneity in residual variance has consequences: one may statistically 'force' between-individual variance to change as a function of the environmental covariate (Fig. 2b) because the error is inappropriately modelled. This may then induce biased estimates or significance of the variance in plasticity ('I × E'). Moreover, failure to detect heterogeneous residuals also precludes any consideration of ecological causes of this type of variation.

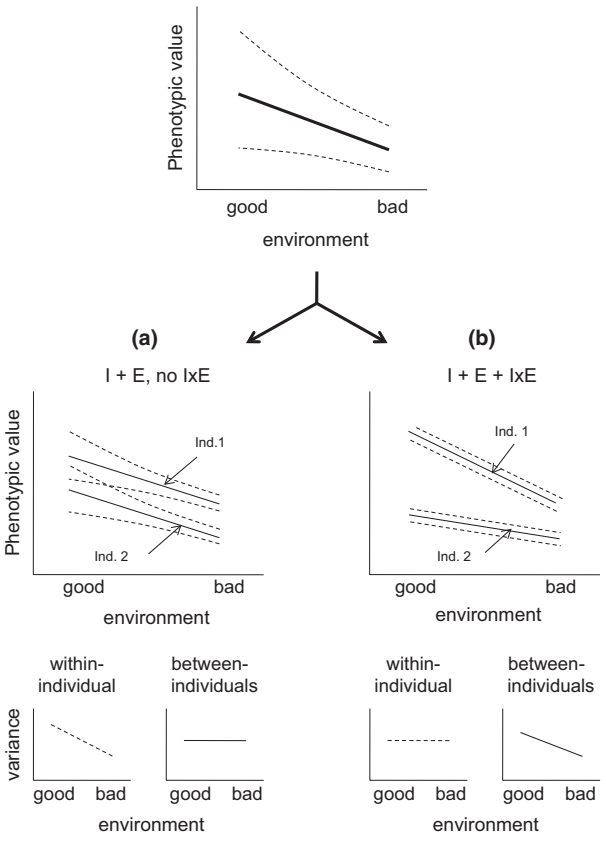

Fig. 2 Conceptual representation of a reaction norm with heterogeneous vs. homogeneous residual variance. Consider a reaction norm where the population-average value of the trait (depicted by the bold line) and its variance (depicted by the dashed lines) decrease as a function of environmental quality. Panels a, b: Random regression models allow partitioning the variance in the reaction norm components into between (solid lines)- and within-individual (dashed line) components (we present here only two individuals in the population). The decrease in trait variance may have two sources: (a) it may be caused by a decrease in within-individual variance (i.e. heterogeneity of residual variance) but not in between-individual variance along the environmental gradient. In this case, individual differences in elevation (I) but not in plasticity (I × E) will be detected; (b) it may be caused by a decrease in between-individual variance but not in within-individual variance (i.e. homogeneity of residual variance) along the environmental gradient. In this case, individual differences in elevation (I) and in plasticity (I × E) will be detected. Because these two sources are not mutually exclusive, a combination of both may drive the change in trait variance (not shown here).

Therefore, to explore whether the pattern of I × E was affected by heterogeneous patterns of within-individual

variance over density gradients (i.e. years; Fig. 2), in both eqns 1 and 2, residual errors $\varepsilon_{i,j,k,l}$ were first modelled as homogeneous (Table S2 in Appendix S2), that is, a common within-individual among-year variance was estimated and assumed to be the same over the entire density gradient. We then assumed that residual variance was instead heterogeneous and allowed it to be specific to each year *j*-level at which population density varies (estimating 18 year-specific residual variances, V_{Rj}). In other words, we considered that observed clutch sizes could deviate more from each individual's estimated reaction norm in certain years or density periods. Notably, an important consequence of having V_{Rj} and consequently the total phenotypic variance specific to each year *j* (V_{Pj}) is that heritability of clutch size also varies as a function of year. For each of the 18 values of annual density, we thus calculated, based on eqn (2), narrow-sense heritability as $(h_j^2) = V_{A0}/V_{Pj}$ where V_{Pj} was calculated as the sum of all variance components included in the most parsimonious model.

Because annual breeding density differed strongly between period 1 ('low'-density years: 1994–2004) and period 2 ('high'-density years: 2005–2011) of the study ($F_{1,18} = 128.26$, $P < 0.001$; Fig. 1a), we also analysed clutch size–density reaction norms in models where residuals errors were allowed to be specific for the 'low'- or 'high'-density period (estimating 2 residual variances) instead. This additional analysis allowed us to investigate whether density-dependent patterns of clutch size were due to a steady change along the density gradient or due to an adjustment to large density changes (Table S3 in Appendix S3). Heritability of clutch size was calculated separately for these two ('low' vs. 'high') density periods.

Results

Clutch size–density reaction norms

Models with heterogeneous errors (i.e. heterogeneous within-individual among-year variance) provided a substantially better fit compared with models with homogenous ones (LRT between model 4 of Table S2 in Appendix S2 vs. model 3 in Table 1: $\chi^2 = 78.46$, d.f. = 17, $P < 0.001$). Therefore, only models with heterogeneous errors will be presented below, but we will highlight any differences in the estimates of random parameters between these two approaches.

Consistent with previous studies of this species (Klomp, 1970; Nicolaus *et al.*, 2009a), clutch size decreased linearly and significantly with both annual breeding density and lay date (quadratic effects of these covariates were not significant, Table 1, Fig. 2b). Clutch

Table 1 Results of the univariate random regression animal model of clutch size as a function of annual breeding density. (a) Estimates of random regression variance components are given with their standard error in parentheses (as specified in eqns 1 and 2). For each model, variance terms are provided with a likelihood ratio test (LRT; χ^2 statistics with associated d.f.) between the given model and the previous model. The LRT was based on a mixture of chi-square probability distributions with 0 and 1 d.f. (indicated by d.f. = 0.5) when testing a single variance component and an equal mixture of chi-square probabilities with 1 and 2 d.f. (indicated by d.f. = 1.5) for tests involving one variance and one covariance. The most parsimonious model is denoted in bold face. All models are fitted with *heterogeneous* residual variance. (b) Estimates of the fixed effects of model 4 (random slopes model) and the most parsimonious model (model 5) (β) are given with their standard error in parentheses. Their significance is tested using *F* tests.

(a)	Random regression variance							Test		
Model	Year	Plot	I(ind_0)	I × E(ind_1)	G(a_0)	PE(pe_0)	LogL	χ^2	d.f.	*P*
1	0.259 (0.100)	–	–	–	–	–	−2267.99	180.82	0.5	< 0.0001
2	0.283 (0.108)	0.216 (0.079)	–	–	–	–	−2198.08	139.82	0.5	< 0.0001
3	0.272 (0.103)	0.198 (0.075)	0.957 (0.075)	–	–	–	−2108.58	179	0.5	< 0.0001
4	0.270 (0.102)	0.194 (0.074)	0.887 (0.096)	0.137 (0.090)	–	–	−2107.24	2.68	1.5	0.182
5	**0.272 (0.103)**	**0.178 (0.069)**	–	–	**0.532 (0.124)**	**0.444 (0.122)**	**−2097.45***	**22.26**	**0.5**	**< 0.0001**

	Fixed effects†					Fixed effects‡				
(b)	β	Wald's *F*	d.f. (nom.)	d.f. (den.)	*P*	β	Wald's *F*	d.f. (nom.)	d.f. (den.)	*P*
Intercept	9.182 (0.325)	799.22	1	277.10	< 0.001	9.167 (0.322)	807.33	1	274.80	< 0.001
Density	−0.784 (0.202)	15.00	1	24.70	< 0.001	−0.779 (0.202)	14.80	1	24.40	< 0.001
Lay date	−0.043 (0.006)	54.83	1	1862.40	< 0.001	−0.042 (0.006)	53.63	1	1866.20	< 0.001
Age	0.506 (0.139)	13.30	1	1203.60	< 0.001	0.503 (0.137)	13.42	1	1191.50	< 0.001
Age^2	−0.063 (0.020)	10.05	1	1067.30	0.002	−0.062 (0.020)	9.88	1	1054.40	0.002

*This LRT is performed by comparing Loglikelihoods between model 3 and 5.

†Fixed effects of model 4.

‡Fixed effects of model 5.

size also exhibited a significant quadratic relationship with female age which is consistent with patterns of senescence patterns reported for this species (Kluijver, 1951; Bouwhuis *et al.*, 2010). Estimates of fixed effects were qualitatively the same in the most parsimonious model and the model fitted with random slopes, implying that our method did not produce any biases (Table 1). The RRAM further revealed that *year* and *plot* explained on average 11% (range: 7–17%; $n = 18$ year-plot specific variances) and 8% (range: 6–13%; $n = 18$ year-plot-specific variances) of the variance in clutch size, respectively, implying that there was considerable spatial and temporal variation that was not attributable to the fixed effects alone (models 1 and 2; Table 1). The inclusion of ind_0, the variance among individuals in reaction norm elevation ('I' for individual), increased the fit of the model significantly (model 3; Table 1), explaining an average 39% of the variance in clutch size (range: 25–59%). This finding implies that females differed significantly in their clutch size at the mean-centred density. Interestingly, the inclusion of ind_1, the between-individual variance in reaction norm slopes (I × E for individual-environment interaction), did not increase the fit of the model significantly when heterogeneous (i.e. year-specific) errors were assumed (model 4; Table 1), whereas it did increase the fit of the model significantly when modelling homogeneous errors (model 5; Table S2 in Appendix S2). This finding implies that all females adjust their clutch size to a change in density in a similar way (i.e. there was no statistical support for individual variation in plasticity) and that the estimate of I × E obtained in models with homogeneous errors (0.171 ± 0.087; model 5 in Table S2 in Appendix S2) was biased upwards (i.e. both in terms of the estimate and value of significance) compared with the estimate of I × E in the model where residual variances were year-specific (0.137 ± 0.090; model 4 in Table 1). Splitting I into its additive genetic (a_0 or 'G') and environmental (pe_0 or 'PE') components further improved the fit of the model (model 5; Table 1), implying that clutch size was heritable in this population. Because there was no statistical evidence for I × E, we logically concluded – following Nussey *et al.* (2007) – that we also could not provide statistical support for genetic variation in slope of clutch size–density reaction norm.

Heritability

Clutch sizes produced by individuals deviated more from their estimated reaction norms in certain years and density periods, and consequently the residual (i.e. within-individual among-year) variance was heterogeneous with respect to these factors: it fluctuated between years of the same density period and decreased from the low- to high-density period (Fig. 3a, Table 2). Consequently, estimates of narrow-sense heritability (h^2) also varied between years of the same density period and increased by 21% between the low- and high-density periods (Fig. 3b, Table 2).

Discussion

Great tit females reduce their clutch size with increasing annual breeding density (Both, 1998a; Nicolaus *et al.*, 2009a). This paper investigated individual variation in such reaction norms and showed that females differed in their expected clutch size when controlling for density (i.e. $I > 0$) and that part of this between-individual variation in clutch size was heritable (cf. Postma & van Noordwijk, 2005). We did not detected significant individual variation in plasticity (i.e. slopes of the clutch size–density reaction norms or I × E) when within-individual among-year or among-period variance in clutch size was modelled as heterogeneous. In contrast, models with homogeneous within-individual variance falsely detected significant I × E.

Between-individual variation

Our study revealed that females differed in elevation of their clutch size–density reaction norms and that the heritability of clutch size in the Lauwersmeer population

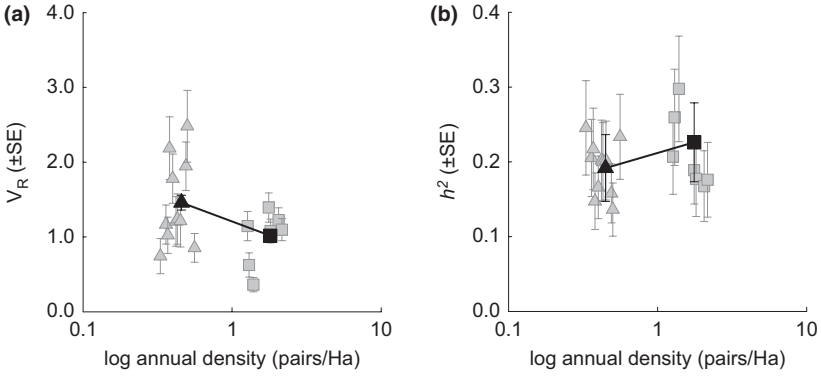

Fig. 3 (a) Mean residual (i.e. within-individual among-year) variance in clutch size ($V_R \pm$ standard error) and (b) mean heritability ($h^2 \pm$ standard error) estimates for all 18 study years (grey) and for the low- and high-density periods (black) as a function of annual breeding density. All estimates are derived from models 3 and 5 of Table 1 and models 5 and 7 of Table S2. 'Low'- and 'high'-density periods are denoted with triangles and squares, respectively.

Table 2 Annual breeding density (pairs per ha), residual variance (V_R) and narrow-sense heritability (h^2) in clutch size specific for (a) each year *j* of the study period (1994–2011) and for (b) each density period *p* (low-density period: 1994–2004; high-density period: 2005–2011). All parameters (±standard error) are obtained from univariate random regression animal models of clutch size as a function of annual breeding density where residual (i.e. within-individual among-year) variance was allowed to vary with annual breeding density (a) or with density periods (b).

(a)

Year	Density	V_{Rj}	h^2_j
1994	0.33	0.744 (0.236)	0.245 (0.063)
1995	0.36	1.167 (0.261)	0.205 (0.052)
1998	0.37	1.024 (0.243)	0.217 (0.054)
1996	0.38	2.186 (0.419)	0.147 (0.038)
1997	0.4	1.781 (0.33)	0.166 (0.042)
2001	0.42	1.208 (0.332)	0.202 (0.054)
2000	0.43	1.24 (0.336)	0.200 (0.053)
2002	0.45	1.213 (0.347)	0.202 (0.053)
1999	0.49	1.946 (0.325)	0.158 (0.040)
2003	0.5	2.48 (0.481)	0.136 (0.036)
2004	0.56	0.854 (0.192)	0.234 (0.057)
2009	1.27	1.145 (0.195)	0.207 (0.050)
2011	1.3	0.625 (0.162)	0.260 (0.064)
2006	1.39	0.363 (0.097)	0.298 (0.071)
2010	1.76	1.394 (0.197)	0.189 (0.045)
2008	1.81	1.079 (0.16)	0.177 (0.050)
2005	2.05	1.226 (0.164)	0.168 (0.047)
2007	2.16	1.097 (0.148)	0.176 (0.050)

(b)

Period	Mean density	V_{Rp}	h^2_p
Low (1994–2004)	0.426 (0.024)	1.458 (0.098)	0.192 (0.045)
High (2005–2011)	1.677 (0.086)	1.017 (0.065)	0.226 (0.053)

varied between 0.14 and 0.30, which is very similar to the estimates reported for other great tit populations where year-specific within-individual variation was not modelled (reviewed by Postma & van Noordwijk, 2005). Proximate explanations for individual differences in elevation may involve nonrandom settlement in heterogeneous environment (Högstedt, 1980) or individual variation in the ability to exploit/monopolize resources or to use and integrate environmental cues (e.g. with females laying larger clutches being better in exploiting heterogeneous resources or being more sensitive to changes in food situation; Mathot *et al.*, 2012). In the first case, the heritability estimate of clutch size may thus partly be explained by permanent environmental effects (e.g. females repeatedly laying larger clutches in good territories). Future studies should aim at modelling heterogeneity in residual variance at the plot or territory level to test whether accounting for limited dispersal or spatial heterogeneity in resources would reduce estimate of heritability (Stopher *et al.*, 2012).

Furthermore, the fact that between-individual differences in plasticity were not significant suggests that in this population, the clutch size–density reaction norm may be under stabilizing selection: From an adaptive viewpoint, any deviation from this reaction norm may have been selected against leading to a decrease in population variance in slopes. If so, the negative density dependence of reproduction, a common phenomenon in passerines, may be interpreted as an adaptive response to an increase in local competition for food prior to, during (e.g. Both, 1998b), or following the nestling phase (Nicolaus *et al.*, 2009a). However, the absence of significant I × E does not preclude the presence of significant selection on phenotypic plasticity (cf. Charmantier *et al.*, 2008), implying that insight into patterns of selection must await their actual quantification. At the same time, individuals may also be genetically and environmentally constrained in differentially adjusting their clutch size to population density (Auld *et al.*, 2010). Alternatively, our measure of annual density may only capture part of the overall underlying process of competition that influence clutch size. If birds tune their clutch size to density at a small spatial scale (i.e. within-plot density or density nearby neighbours) rather than to density of the population as a whole, lack of significant I × E might be the result. This explanation is unlikely because we already showed elsewhere that annual – not local plot – density drives density dependence of clutch size (Nicolaus *et al.*, 2009a). Nevertheless, because annual breeding densities increased with year in period 1 and increased between periods (Fig. 1a), there remains the possibility that females tune their clutch size in earlier years to some environmental parameter(s) other than density that changed in correspondence with period (discussed below). Finally, we may not have been able to detect significant I × E, because for females with < 3 observations, statistical power to detect I × E was around 0.6 (although power ranged between 0.6 and 1.0 for the data set as a whole) and imprecision around the estimate was relatively large (see Appendix S1).

Environment-specific within-individual variation

Observed clutch sizes deviated much more from the individual's estimated reaction norm in certain years compared with others, implying within-individual among-year heterogeneous patterns of residual variance. When plotted against density, the pattern reveals that changes in within-individual among-year deviations from reaction norms were not related to subtle year-to-year changes in density (e.g. no 'fanning in' pattern as hypothesized in Fig. 2a). This suggests that changes in within-individual among-year variance (Fig. 3a) could equally well be explained by fluctuations in other environmental factors than density (e.g. changes in food availability or other resources such as

nest boxes). Several proximate mechanisms may explain why in certain years individuals produced clutch sizes that deviate more strongly from their overall clutch size–density reaction norm. First, it may be linked to annual variation in the ecological conditions and to the precision of environmental cues that females rely on when taking reproductive decisions (DeWitt, 1998; Auld *et al.*, 2010). When cues do not reflect reliably the environmental conditions, individuals may be less likely to produce the optimal phenotype that matches the specific environment. It could then be more advantageous to increase the variance in clutch size as cues become increasingly flawed (DeWitt & Langerhans, 2004). Second, clutch size may be part of a multidimensional reaction norm (Westneat *et al.*, 2009, 2011). Because birds live in environments that vary in many ways, their clutch size may be function of multiple environmental factors. Heterogeneous errors may, in this case, reflect interactions or correlations between density and other variables that may vary more in some years than others (Westneat *et al.*, 2009). Third, some years may be associated with lower food availability that may reduce female endogenous reserve accumulation prior to breeding and homogenize variability in condition among females (Arcese & Smith, 1988; Nager *et al.*, 1997). As a consequence, energy constraints on egg production may be greater, and females may be forced to reduce their reproductive investment, thereby reducing within-individual among-year variance (Arcese & Smith, 1988; Nager *et al.*, 1997). Condition dependence in reproductive investment may be a general mechanism explaining limitations in a trait expression (Nussey *et al.*, 2005a; Reed *et al.*, 2009). Heterogeneity in residuals may thus reflect biologically relevant mechanisms, and future studies should aim at distinguishing between these (nonexhaustive) options. Specifically, investigation of patterns of residual variance will be particularly important with respect to the environmental gradient under study (in our case: yearly density changes) because doing so will help avoid erroneous conclusions. For example, had we focussed on spatial variation in density instead, analyses of heterogeneity at the level of plot rather than year would have been warranted.

Our results also show that the within-individual among-year variance was larger for individuals breeding under low compared with high-density periods (as depicted in Fig. S4 in Appendix S4). The experimental increase in density between the two time periods may have caused the ecological conditions to deteriorate (e.g. via an increase of intraspecific competition, a change in the social composition of the population or a change in individual territory quality). In this case, selection may have canalized the level of expressed variation of clutch size within individuals under harsh environmental conditions (Fig. S4 in Appendix S4; Gibson & Wagner, 2000). This hypothesis now warrants experimental verification. Such experiments should logically include manipulations of yearly density and thereby reveal whether the presumed effects of density revealed by this observational study were indeed causal rather than related to other correlated environmental factors.

Heterogeneity in residual variances: methodological implications

Our study shows that allowing residual variances to vary with annual density improved the model fit and reduced the estimate of I × E [homogeneous (SE) vs. heterogeneous (SE) = 0.171 (0.089) vs. 0.137 (0.090)]. This effect of heterogeneous residual variance on the estimation of random regression parameters is important for analyses that use a frequentist's approach (i.e. where inferences are based on values of p) because it affects the interpretation of the results. For example, with homogeneous residual variances (model 5 of Table S2 in Appendix S2), we would have concluded that individuals differ significantly in plasticity of clutch size, which was not the case with heterogeneous residual variances. Ignoring this heterogeneity inflated between-individual variance in plasticity by a value of almost 20% because it forced the between-individual variance to vary with density. In general, the problem of heterogeneous errors should be analysed for data presenting spatial or temporal components: in natural populations, repeated measures of individuals often tend to be clustered in space or time, and therefore, consecutive measurements within an individual are likely to be collected under similar conditions (van der Jeugd & McCleery, 2002). Because the random regression procedure estimates individual slopes over the continuum of the covariate, these differences in within-individual variance will become associated with between-individual differences. Consequently, these within-individual differences can be erroneously interpreted as between-individual variances (Fig. 2b) and thus cause bias in the estimated variance components.

Based on published literature, it appears that most empirical studies in behavioural ecology ignore heterogeneity in residual variance, particularly those focussing on behaviour (11 of 12 studies (91.7%) in Table 3; see also Cleasby & Nakagawa, 2011); this problem applies to a lesser extent also to studies of life-history traits (four of 13 studies (30.8%) in Table 3). Therefore, the amount of I × E may be generally overestimated. Future studies should investigate patterns in residual variances, especially because those studies that considered both homogeneous vs. heterogeneous within-individuals variances typically found that the latter fitted the data best (Brommer *et al.*, 2008; Reed *et al.*, 2009; Husby *et al.*, 2010; Dingemanse *et al.*, 2012a; Westneat *et al.*, 2013).

Table 3 Studies that have used random regression analysis to estimate individual variation in plasticity (I × E) in behavioural (B) or life-history (LH) traits. Statistical evidence (i.e. $P < 0.05$) for I × E or heterogeneity in residual (i.e. within-individual) variances are indicated by yes (Y) or no (N). The list of studies was adapted and modified from the literature reviews by Mathot *et al.* (2012) and van de Pol (2012).

Number	Species	Trait type	Trait	Environmental covariate	Evidence for I × E?	Allowing heterogeneity in residual variances?	Reference
1	Brown trout (*Salmo trutta*)	B	Aggressiveness/ Foraging activity	Experience	N/Y	N	Adriaenssens & Johnsson (2011)
2	Eastern chipmunk (*Tamias striatus*)	B	Activity/Exploration	Familiarity	N	N	Martin & Réale (2008)
3	Great tit (*Parus major*)	B	Exploration	Familiarity/ Season/Interval	Y	N	Dingemanse *et al.* (2012b)
4	Great tit (*Parus major*)	B	Feeding behaviour	Predation risk	Y	N	Quinn *et al.* (2012)
5	House sparrow (*Passer domesticus*)	B	Provisioning rate	Nestling age/partner	Y	N	Westneat *et al.* (2011)
6	Nutmeg mannikins (*Lonchura punctulata*)	B	Feeding rate/vigilance	Group size	N/N	N	Rieucau *et al.* (2010)
7	Nutmeg mannikins (*Lonchura punctulata*)	B	Patch use/ scrounger use	Food	Y/Y	N	Morand Ferron *et al.* (2011)
8	Red knots (*Calidris canutus*)	B	Escape flight duration/ Proportion of time vigilant	Predation danger	N/Y	N	Mathot *et al.* (2011)
9	Speckled damselfish (*Pomacentrus bankanensis*)	B	Activity/Boldness	Temperature	Y/Y	N	Biro *et al.* (2010)
10	Yellow damselfish (*Pomacentrus miluccensis*)	B	Activity/Aggression/ Boldness	Temperature	Y/N/N	N	Biro *et al.* (2010)
11	Threespined stickleback (*Gasterosteus aculeatus*)	B	Exploration	Familiarity Predation risk	Y	Y	Dingemanse *et al.* (2012a)
12	Tree swallow (*Tachycineta bicolor*)	B	Aggression	Temperature	Y	N	Betini & Norris (2012)
13	Ural owl (*Strix uralensis*)	B	Aggression	Food availability	Y	N	Kontiainen *et al.* (2009)
14	Collared flycatcher (*Ficedula albicollis*)	LH	Laying date	3 weather variables	Y	N	Brommer *et al.* (2005)
15	Common gull (*Larus canus*)	LH	Laying date	Spring temperature	Y	Y	Brommer *et al.* (2008)
16	Eurasian oystercatcher (*Haematopus ostralegus*)	LH	Laying date	Age	N	N	van de Pol and Verhulst (2006)
17	Great tit (*Parus major*)	LH	Laying date	Spring temperature	N	N	Charmantier *et al.* (2008)
18	Great tit (*Parus major*)	LH	Laying date	Spring temperature	Y	N	Nussey *et al.* (2005a)
19	Great tit (*Parus major*)	LH	Laying date	Spring temperature	Y	Y	Husby *et al.* (2010)
20	Great tit (*Parus major*)	LH	Clutch size	Annual breeding density	N	Y	This study
21	Guillemot (*Uria aalge*)	LH	Laying date	Various	Y	Y	Reed *et al.* (2009)
22	Guillemot (*Uria aalge*)	LH	Laying date	NAO	N	N	Reed *et al.* (2006)
23	House sparrow (*Passer domesticus*)	LH	Clutch size	Age/laying date	N	N	Westneat *et al.* (2009)
24	Painted turtle (*Chrysemys picta*)	LH	Laying date	Winter temperature	Y	N	Schwanz & Janzen (2008)
25	Red deer (*Cervus elaphus*)	LH	Calving date	Rainfall	Y	N	Nussey *et al.* (2005b)
26	Ural owl (*Strix uralensis*)	LH	Clutch size	Laying date	Y	N	Brommer *et al.* (2012)

Concluding remarks

Our study highlights the importance of using long-term individual-based data to understand patterns of phenotypic variation. By quantifying variation in plasticity at the individual level, we showed that plasticity at the population level does not necessarily imply that variation in plasticity exists at the level of the individual (no I × E or G × E). We have shown that within-individual among-year deviations from reaction norms were not constant over the years or density periods and that this influenced the heritability of clutch size. Although variation in heritability was small in our study, this finding supports accumulating evidence that heritability is not only specific to a given population but can vary also across environments within populations (e.g. Husby *et al.*, 2010, 2011). The reported patterns thereby open up interesting future questions about how varying ecological conditions can affect evolutionary trajectories of phenotypic traits. Overall, this study confirms that (behavioural) reaction norm approaches are powerful tools to investigate the structure of phenotypic variance (Nussey *et al.*, 2007; Dingemanse *et al.*, 2010; Westneat *et al.*, 2011; Dingemanse & Dochtermann, 2013). This approach may also shed light on the ultimate and proximate factors influencing the complex nature of this variation (Dingemanse & Wolf, 2013). Our study thereby calls for studies considering the biology of heterogeneity in residual errors.

Acknowledgments

We thank all the cohorts of people who have helped collecting great tit breeding data over the last 18 years. We also thank Staatsbosbeheer and the Royal Dutch Army 'Koninklijke Landmacht' for permission to work on their properties. We are grateful to Dave Westneat, Dan Nussey and an anonymous referee for providing constructive comments on the manuscript and to Dirk Metzel for providing statistical advice. M.N. was supported by the Alexander von Humboldt Foundation and by the Max Planck Society (MPG), J.E.B by the Academy of Finland, R.U. and J.M.T by the University of Groningen and N.J.D. by the Max Planck Society (MPG).

References

Adriaenssens, B. & Johnsson, J.I. 2011. Shy trout grow faster: exploring links between personality and fitness-related traits in the wild. *Behav. Ecol.* **22**: 135–143.

Arcese, P. & Smith, J.N.M. 1988. Effects of population density and supplemental food on reproduction in Song Sparrows. *J. Anim. Ecol.* **57**: 119–136.

Auld, J.R., Agrawal, A.A. & Relyea, R.A. 2010. Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proc. Biol. Sci.* **277**: 503–511.

Betini, G.S. & Norris, D.R. 2012. The relationship between personality and plasticity in tree swallow aggression and the consequences for reproductive success. *Anim. Behav.* **83**: 137–143.

Biro, P.A., Beckmann, C. & Stamps, J.A. 2010. Small within-day increases in temperature affects boldness and alters personality in coral reef fish. *Proc. Biol. Sci.* **277**: 71–77.

Both, C. 1998a. Density dependence of clutch size: habitat heterogeneity or individual adjustment? *J. Anim. Ecol.* **67**: 659–666.

Both, C. 1998b. Experimental evidence for density dependence of reproduction in great tits. *J. Anim. Ecol.* **67**: 667–674.

Both, C. 2000. Density dependence of avian clutch size in resident and migrant species: is there a constraint on the predictability of competitor density? *J. Avian Biol.* **31**: 412–417.

Both, C., Tinbergen, J.M. & Visser, M.E. 2000. Adaptive density dependence of avian clutch size. *Ecology* **81**: 3391–3403.

Bouwhuis, S., Charmantier, A., Verhulst, S. & Sheldon, B. 2010. Trans-generational effects on ageing in a wild bird population. *J. Evol. Biol.* **23**: 636–642.

Brommer, J.E. in press. Phenotypic plasticity of labile traits in the wild. *Curr. Zool.*

Brommer, J.E., Merilä, J., Sheldon, B.C. & Gustasfsson, L. 2005. Natural selection and genetic variation for reproductive reaction norms in a wild bird population. *Evolution* **59**: 1362–1371.

Brommer, J.E., Rattiste, K. & Wilson, A.J. 2008. Exploring plasticity in the wild: laying date-temperature reaction norms in the common gull *Larus canus*. *Proc. Biol. Sci.* **275**: 687–693.

Brommer, J.E., Rattiste, K. & Wilson, A. 2010a. The rate of ageing in a long-lived bird is not heritable. *Heredity* **104**: 363–370.

Brommer, J.E., Alho, J.S., Biard, C., Chapman, J.R., Charmantier, A., Dreiss, A. *et al.* 2010b. Passerine extrapair mating dynamics: A bayesian modeling approach comparing four species. *Am. Nat.* **176**: 178–187.

Brommer, J.E., Kontiainen, P. & Pietiäinen, H. 2012. Selection on plasticity of seasonal life-history traits using random regression mixed model analysis. *Ecol. Evol.* **2**: 695–704.

Charmantier, A. & Réale, D. 2005. How do misassigned paternities affect the estimation of heritability in the wild? *Mol. Ecol.* **14**: 2839–2850.

Charmantier, A., McCleery, R., Cole, L.R., Perrins, C.M., Kruuk, L.E.B. & Sheldon, B.C. 2008. Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* **320**: 800–803.

Cleasby, I.R. & Nakagawa, S. 2011. Neglected biological patterns in the residuals A behavioural ecologist's guide to cooperating with heteroscedasticity. *Behav. Ecol. Sociobiol.* **65**: 2361–2372.

DeWitt, T.J. 1998. Costs and limits of phenotypic plasticity: tests with predator-induced morphology and life history in a freshwater snail. *J. Evol. Biol.* **11**: 465–480.

DeWitt, T.J. & Langerhans, R.B. 2004. Integrated solutions to environmental heterogeneity: theory of multimoment reaction norms. In: *Phenotypic Plasticity: Functional and Conceptual Approaches* (T.J. DeWitt & S.M. Scheiner, eds), pp. 98–111. University of Oxford, Oxford.

DeWitt, T.J. & Scheiner, S.M. 1994. *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford University Press, Oxford.

Dingemanse, N.J. & Dochtermann, N.A. 2013. Quantifying individual variation in behaviour: mixed-effect modelling approaches. *J. Anim. Ecol.* **82**: 39–54.

Dingemanse, N.J. & Wolf, M. 2013. Between-individual differences in behavioural plasticity within populations: causes and consequences. *Anim. Behav.* **85**: 1031–1039.

Dingemanse, N.J., Kazem, A.J.N., Réale, D. & Wright, J. 2010. Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol. Evol.* **25**: 81–89.

Dingemanse, N.J., Barber, I., Wright, J. & Brommer, J.E. 2012a. Quantitative genetics of behavioural reaction norms: genetic correlations between personality and behavioural plasticity vary across stickleback populations. *J. Evol. Biol.* **25**: 485–496.

Dingemanse, N.J., Bouwman, K.M., van de Pol, M., van Overveld, T., Patrick, S.C., Matthysen, E. *et al.* 2012b. Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. *J. Anim. Ecol.* **81**: 116–126.

Erceg-Hurn, D.M. & Mirosevich, V.M. 2008. Modern robust statistical methods an easy way to maximize the accuracy and power of your research. *Am. Psychol.* **63**: 591–601.

Gibson, G. & Wagner, G. 2000. Canalization in evolutionary genetics: a stabilizing theory? *BioEssays* **22**: 372–380.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. & Thompson, R. 2009. *ASReml User Guide Release 3.0*. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Henderson, C.R. 1982. Analysis of covariance in the mixed model – higher-level, non-homogeneous, and random regressions. *Biometrics* **38**: 623–640.

Högstedt, G. 1980. Evolution of clutch size in birds. Adaptive variation in relation to territory quality. *Science* **210**: 1148–1150.

Husby, A., Nussey, D.H., Visser, M.E., Wilson, A.J., Sheldon, B.C. & Kruuk, L.E.B. 2010. Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (*Parus major*) populations. *Evolution* **64**: 2221–2237.

Husby, A., Visser, M.E. & Kruuk, L.E.B. 2011. Speeding up microevolution: the effects of increasing temperature on selection and genetic variance in a wild bird population. *PLoS Biol.* **9**: e1000585.

van der Jeugd, H.P. & McCleery, R. 2002. Effects of spatial autocorrelation, natal philopatry and phenotypic plasticity on the heritability of laying date. *J. Evol. Biol.* **15**: 380–387.

Klomp, H. 1970. Determination of clutch size in birds. A review. *Ardea* **58**: 1–124.

Kluijver, H.N. 1951. The population ecology of the Great Tit, *Parus m. major* L. *Ardea* **39**: 1–135.

Kontiainen, P., Pietiäinen, H., Huttunen, K., Karell, P., Kolunen, H. & Brommer, J.E. 2009. Aggressive Ural owl mothers recruit more offspring. *Behav. Ecol.* **20**: 789–796.

Kruuk, L.E.B. 2004. Estimating genetic parameters in natural populations using the 'animal model'. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **359**: 873–890.

Lima, S.L. 2009. Predators and the breeding bird: behavioral and reproductive flexibility under the risk of predation. *Biol. Rev.* **84**: 485–513.

Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderlands, MA, USA.

Martin, J.G.A. & Réale, D. 2008. Temperament, risk assessment and habituation to novelty in eastern chipmunks, *Tamias striatus*. *Anim. Behav.* **75**: 309–318.

Martin, J.G.A., Nussey, D.H., Wilson, A.J. & Réale, D. 2011. Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. *Methods Ecol. Evol.* **2**: 362–374.

Mathot, K.J., van den Hout, P.J., Piersma, T., Kempenaers, B., Réale, D. & Dingemanse, N.J. 2011. Disentangling the roles of frequency-vs. state-dependence in generating individual differences in behavioural plasticity. *Ecol. Lett.* **14**: 1254–1262.

Mathot, K.J., Wright, J., Kempenaers, B. & Dingemanse, N.J. 2012. Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. *Oikos* **121**: 1009–1020.

Meyer, K. 1998. Estimating covariance functions for longitudinal data using a random regression model. *Genet. Sel. Evol.* **30**: 221–240.

Morand-Ferron, J., Wu, G.M. & Giraldeau, L.A. 2011. Persistent individual differences in tactic use in a producer-scrounger game are group dependent. *Anim. Behav.* **82**: 811–816.

Morrissey, M.B. & Wilson, A.J. 2010. Pedantics: an r package for pedigree-based genetic simulation and pedigree manipulation, characterization and viewing. *Mol. Ecol. Resour.* **10**: 711–719.

Nager, R.G., Ruegger, C. & van Noordwijk, A.J. 1997. Nutrient or energy limitation on egg formation: a feeding experiment in great tits. *J. Anim. Ecol.* **66**: 495–507.

Newton, I. 1998. *Population Limitation in Birds*. Academic Press, London.

Nicolaus, M., Both, C., Ubels, R., Edelaar, P. & Tinbergen, J.M. 2009a. No experimental evidence for local competition in the nestling phase as a driving force for density-dependent avian clutch size. *J. Anim. Ecol.* **78**: 828–838.

Nicolaus, M., Michler, S.P.M., Ubels, R., van der Velde, M., Komdeur, J., Both, C. *et al.* 2009b. Sex-specific effects of altered competition on nestling growth and survival: an experimental manipulation of brood size and sex ratio. *J. Anim. Ecol.* **78**: 414–426.

Nussey, D.H., Clutton-Brock, T.H., Elston, D.A., Albon, S.D. & Kruuk, L.E.B. 2005a. Phenotypic plasticity in a maternal trait in red deer. *J. Anim. Ecol.* **74**: 387–396.

Nussey, D.H., Postma, E., Gienapp, P. & Visser, M.E. 2005b. Selection on heritable phenotypic plasticity in a wild bird population. *Science* **310**: 304–306.

Nussey, D.H., Wilson, A.J. & Brommer, J.E. 2007. The evolutionary ecology of individual phenotypic plasticity in wild populations. *J. Evol. Biol.* **20**: 831–844.

Piersma, T. & Drent, J. 2003. Phenotypic flexibility and the evolution of organismal design. *Trends Ecol. Evol.* **18**: 228–233.

Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going now? *Trends Ecol. Evol.* **20**: 481–486.

Pinheiro, J.C. & Bates, D.M. 2000. *Mixed Effect Models in S and S-PLUS*. Springer, New York, NY, USA.

van de Pol, M. 2012. Quantifying individual variation in reaction norms: how study design affects the accuracy, precision and power of random regression models. *Methods Ecol. Evol.* **3**: 268–280.

van de Pol, M. & Verhulst, S. 2006. Age-dependent traits: a new statistical model to separate within-and between-individual effects. *Am. Nat.* **167**: 766–773.

Postma, E. & van Noordwijk, A.J. 2005. Genetic variation for clutch size in natural populations of birds from a reaction norm perspective. *Ecology* **86**: 2344–2357.

Quinn, J.L., Cole, E.F., Bates, J., Payne, R.W. & Cresswell, W. 2012. Personality predicts individual responsiveness to the risks of starvation and predation. *Proc. Biol. Sci.* **279**: 1919–1926.

Réale, D. & Dingemanse, N.J. 2010. Personality and individual social specialisation. In: *Social Behaviour: Genes, Ecology and Evolution* (T. Székely, A. Moore & J. Komdeur, eds), pp. 417. Cambridge University Press, Cambridge.

Reed, T.E., Wanless, S., Harris, M.P., Frederiksen, M., Kruuk, L.E. & Cunningham, E.J. 2006. Responding to environmental change: plastic responses vary little in a synchronous breeder. *Proc. Biol. Sci.* **273**: 2713–2719.

Reed, T.E., Warzybok, P., Wilson, A.J., Bradley, R.W., Wanless, S. & Sydeman, W.J. 2009. Timing is everything: flexible phenology and shifting selection in a colonial seabird. *J. Anim. Ecol.* **78**: 376–387.

Rieucau, G., Morand-Ferron, J. & Giraldeau, L.A. 2010. Group size effect in nutmeg mannikin: between-individuals behavioral differences but same plasticity. *Behav. Ecol.* **21**: 684–689.

Schaeffer, L.R. 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* **86**: 35–45.

Schielzeth, H. & Forstmeier, W. 2009. Conclusions beyond support: overconfident estimates in mixed models. *Behav. Ecol.* **20**: 416–420.

Schwanz, L.E. & Janzen, F.J. 2008. Climate change and temperature-dependent sex determination: can individual plasticity in nesting phenology prevent extreme sex ratios? *Physiol. Biochem. Zool.* **81**: 826–834.

Self, S.G. & Liang, K.Y. 1987. Large sample properties of the maximum likelihood estimator and the likelihood ratio test on the boundary of the parameter space. *J. Am. Stat. Assoc.* **82**: 605–611.

Shaw, R.G. 1991. The comparison of quantitative genetic parameters between populations. *Evolution* **45**: 143–151.

Stamps, J.A., Briffa, M. & Biro, P.A. 2012. Unpredictable animals: individual differences in intraindividual variability (IIV). *Anim. Behav.* **83**: 1325–1334.

Stopher, K.V., Walling, C.A., Morris, A., Guinness, F.E., Clutton-Brock, T.H., Pemberton, J.M. *et al.* 2012. Shared spatial effects on quantitative genetic parameters: accounting for spatial autocorrelation and home range overlap reduces estimates of heritability in wild red deer. *Evolution* **66**: 2411–2426.

Svensson, L. 1992. *Identification Guide to European Passerines*. BTO, Norfolk.

Tinbergen, J.M. & Sanz J. 2004. Strong evidence for selection for larger brood size in a great tit population. *Behav. Ecol.* **15**: 525–533.

Via, S., Gomulkiewicz, R., Dejong, G., Scheiner, S.M., Schlichting, C.D. & Vantienderen, P.H. 1995. Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol. Evol.* **10**: 212–217.

Visscher, P.M. 2006. A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Res. Hum. Genet.* **9**: 490–495.

Westneat, D.F., Stewart, I.R. & Hatch, M.I. 2009. Complex interactions among temporal variables affect the plasticity of clutch size in a multi-brooded bird. *Ecology* **90**: 1162–1174.

Westneat, D.F., Hatch, M.I., Wetzel, D.P. & Ensminger, A.L. 2011. Individual variation in parental care reaction norms: integration of personality and plasticity. *Am. Nat.* **178**: 652–667.

Westneat, D.F., Schofield, M. & Wright, J. 2013. Parental behavior exhibits among-individual variance, plasticity, and heterogeneous residual variance. *Behav. Ecol.* **24**: 598–604.

Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A. *et al.* 2010. An Ecologist's guide to the animal model. *J. Anim. Ecol.* **79**: 13–26.

Wolf, M., van Doorn, G.S. & Weissing, F.J. 2008. Evolutionary emergence of responsive and unresponsive personalities. *Proc. Natl. Acad. Sci. USA* **105**: 15825–15830.

Wolf, M., Van Doorn, G. & Weissing, F.J. 2011. On the coevolution of social responsiveness and behavioural consistency. *Proc. Biol. Sci.* **278**: 440–448.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Reaction norm power analysis.

Appendix S2 Clutch size-density reaction norms with homogeneous errors.

Appendix S3 Clutch size-density reaction norms with heterogeneous errors fitted for the low and high density period.

Appendix S4 Conceptual representation of the reaction norms clutch size-density reaction norms in the Lauwersmeer population.

Received 29 January 2013; revised 27 May 2013; accepted 30 May 2013